Environmental Research 166 (2018) 707–712

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Google search patterns monitoring the daily health impact of heatwaves in England: How do the findings compare to established syndromic surveillance systems from 2013 to 2017?

Helen K. Green[a], Obaghe Edeghere[a], Alex J. Elliot[a], Ingemar J. Cox[b,c], Roger Morbey[a], Richard Pebody[d], Angie Bone[e], Rachel A. McKendry[f], Gillian E. Smith[a,*]

[a] Real-time Syndromic Surveillance Team, Public Health England, Birmingham, United Kingdom
[b] Department of Computer Science, University College London, London, United Kingdom
[c] Department of Computer Science, University of Copenhagen, Copenhagen, Denmark
[d] Respiratory Diseases Department, Public Health England, London, United Kingdom
[e] Extreme Events, Public Health England, London, United Kingdom
[f] London Centre for Nanotechnology and Division of Medicine, University College London, London, United Kingdom

ARTICLE INFO

Keywords:
Temperature
Heat-related illness
Heat-waves
Syndromic surveillance
Public health

ABSTRACT

One of the implications of climate change is a predicted increase in frequent and severe heatwaves. The impact of heatwaves on the health of the population is captured through real-time syndromic healthcare surveillance systems monitored daily in England during the summer months. Internet search data could potentially provide improved timeliness and help to assess the wider population health impact of heat by capturing a population sub-group who are symptomatic but do not seek healthcare. A retrospective observational study was carried out from June 2013 to September 2017 in England to compare daily trends in validated syndromic surveillance heat-related morbidity indicators against symptom-based heatwave related Google search terms. The degree of correlation was determined with Spearman correlation coefficients and lag assessment was carried out to determine timeliness. Daily increases in frequency in Google search terms during heatwave events correlated well with validated syndromic indicators. Correlation coefficients between search term frequency and syndromic indicators from 2013 to 2017 were highest with the telehealth service NHS 111 (range of 0.684–0.900 by search term). Lag analysis revealed a similar timeliness between the data sources, suggesting Google data did not provide a delayed or earlier signal in the context of England's syndromic surveillance systems. This work highlights the potential benefits for countries which lack established public health surveillance systems to monitor heat-related morbidity and the use of internet search data to assess the wider population health impact of exposure to heat.

1. Background

Heatwaves can significantly impact on the health of the population, with outcomes ranging from dehydration and sunburn through to heatstroke and death. The highest vulnerability is typically seen in the elderly (Benmarhnia et al., 2015), children (Li et al., 2015), and in those with pre-existing medical conditions (Bouchama et al., 2007), although all age groups may be affected, particularly in hot countries, where people are working outdoors without adequate protection (Nelson et al., 2011) or cannot seek shelter (e.g. homeless, Ramin, 2009).

The Heatwave Plan for England was launched in 2004 (PHE, 2015) in response to the 2003 European heatwave which resulted in an estimated 70,000 deaths across Europe (Robine et al., 2008). As part of this Plan, from 1st June to 15th September each year, Public Health England (PHE) routinely monitors on a daily basis heat-related morbidity across a suite of syndromic surveillance systems (Triple S Project, 2011; PHE, 2017). Syndromic surveillance is the near real-time collection, analysis, interpretation and dissemination of health-related data to provide an early warning of threats requiring public health action, or conversely provide reassurance about the absence of impact of such threats. Indicators monitored within these systems in England have been demonstrated to provide a sensitive measurement of the health impact of heat (Smith et al., 2016a, 2016b; Elliot et al., 2014) and

* Correspondence to: Real-time Syndromic Surveillance Team, National Infection Service, Public Health England, 5 St Philip's Place, Birmingham B3 2PW, United Kingdom.
E-mail address: gillian.smith@phe.gov.uk (G.E. Smith).

https://doi.org/10.1016/j.envres.2018.04.002
Received 30 October 2017; Received in revised form 31 March 2018; Accepted 3 April 2018
Available online 28 June 2018
0013-9351/

information can be used to facilitate optimal public health response during the heatwave through ensuring appropriate messaging (Josseran et al., 2010; Pascal et al., 2012; Lall et al., 2017).

The use of novel data sources in health surveillance, such as social media and internet search queries is gaining increasing interest (Nuti et al., 2014). While syndromic surveillance provides a broader measure of illness through a clinical diagnosis (as opposed to laboratory confirmation, which is a more specific end-point for infectious diseases), current syndromic surveillance systems in use in England still rely on an individual consulting healthcare. The potential advantages of online user-generated content include the provision of more timely information and identification of illness in symptomatic individuals, including those who do not seek healthcare. For data sources such as Google Trends (Google, 2017), this comes at a minimal cost and may therefore be useful in countries with sufficient internet coverage but without an established public health surveillance infrastructure. However, there are notable limitations with these data sources, including uncertainty of the underlying reason for an individual searching and of the characteristics of the person searching.

The potential utility of internet search data to monitor heat-related morbidity was demonstrated in Shanghai (Li et al., 2016), with strong correlation between internet searches for *heat stroke* and heatstroke deaths and hospitalised cases. However, availability of health outcome data was not as timely as in England and it is not known how internet search data for heat-related illness compares to outputs from real-time syndromic surveillance systems. During 2017 in England, there was a Level 2 heatwave alert issued (defined as a risk of reaching high temperatures) from 15 to 21st of June, with areas reaching high temperatures (a Level 3 heatwave alert) across most of the country from 17 to 20th June (Met Office, 2017). This article outlines the findings from a retrospective observational study comparing trends in heat-related syndromic surveillance indicators against selected heatwave related internet search terms during this period and for previous heatwaves. This information assists in determining the utility and timeliness of internet search data as a means of assessing population health impacts from heatwaves.

2. Methods

2.1. Data

National data for England for each of the Google search terms and heat-related syndromic indicators was extracted on a daily timescale from 1 June 2013 to 15 September 2017.

Anonymised daily syndromic data for heat impact indicators (Smith et al., 2016a) was accessed from the four national PHE syndromic surveillance systems; NHS 111 (a telehealth service available to the population of England), an in-hours general practitioner (GP) system (GPIH; denominator population of 35 million registered patients across England), an out of hours GP system (GPOOH; approximately 70% of out of hours activity across England), and a sentinel emergency department (ED) surveillance system (35 sentinel EDs across England and Northern Ireland) as previously described (Smith et al., 2016a, 2016b; Harcourt et al., 2017, PHE 2017).

Google search volume data was available from the Google Health Trends Application Programming Interface (API), a tool to explore Google search data similar to the publicly available Google Trends website (Google, 2017). Data is taken from a uniformly distributed random sample of 10–15% of Google web searches updated daily. Daily probabilities were defined as the probability of searching for a specified term on a given day in England, multiplied by ten million to be human readable.

Most disease surveillance studies looking at Google Trends data typically focus on one or two relevant search terms. We sought to expand on this by utilising the clinical codes underlying the heat-related syndromic morbidity indicators to develop a list of search terms that people may use to search Google. Relevant search terms were identified through the following process:

Table 1
Selection of Google search terms.

Terms identified from syndromic surveillance	Additional terms/comments from Google Trends
Heat exhaustion	Heat stroke
Heat rash	Prickly heat
Heatstroke	Heat stroke
	Sunstroke
Heat stroke	No additional relevant terms identified
Heatwave	No additional relevant terms identified
Prickly heat	Heat rash
Sunburn	No additional relevant terms identified
Sunstroke	No additional relevant terms identified
Sun stroke	Heat stroke
	Sunstroke
Heat syncope	Not enough searches carried out in Google to provide data
Heat fatigue	Not enough searches carried out in Google to provide data
Heat oedema	Not enough searches carried out in Google to provide data
Heat prostration	Not enough searches carried out in Google to provide data

1. Based on experience with clinical codes used for existing syndromic surveillance systems, descriptions of such codes corresponding to heat-related syndromic morbidity indicators (e.g. *heatstroke*) were collated and synthesised into a list of code terms (Table 1).
2. Each code term was entered into Google Trends and the related Google internet search terms retained if also searched for at least half as often (Google, 2017) to identify further terms used (e.g. Google users searching for *heatstroke* also searching for *sunstroke*, Table 1) and associated search terms (e.g. *symptoms*). This cut off was used based on the observed distribution of frequency of related search terms and the relevance of these terms.
3. Where multiple variations of the search terms were available (e.g. *heatstroke*, *heat stroke*), there was generally one that had a markedly larger search probability, and this was retained.

The final list of search terms was as follows:

a) Heat exhaustion
b) Heat rash
c) Heat stroke
d) Heatwave
e) Prickly heat
f) Sunburn
g) Sunstroke

2.2. Analysis

Daily time series plots of national syndromic surveillance and Google search data were produced to visually compare patterns from 2013 to 2017 using daily data and seven day moving averages to minimise the impact of day of the week effects (Buckingham-Jeffery et al., 2017). Correlation was quantified through Spearman correlation coefficients as the data was not normally distributed. This comparison was also separately done for 2017 from 8 June 2017 to 27 June 2017 to focus on the known heatwave event, encompassing one week either side of the official heatwave alerts. Lagged time series plots were produced for the 2017 heatwave to assess the time lag/lead at which the two datasets best correlated. The Google time series was shifted forwards or backwards in time up to ten days to determine if search activity was

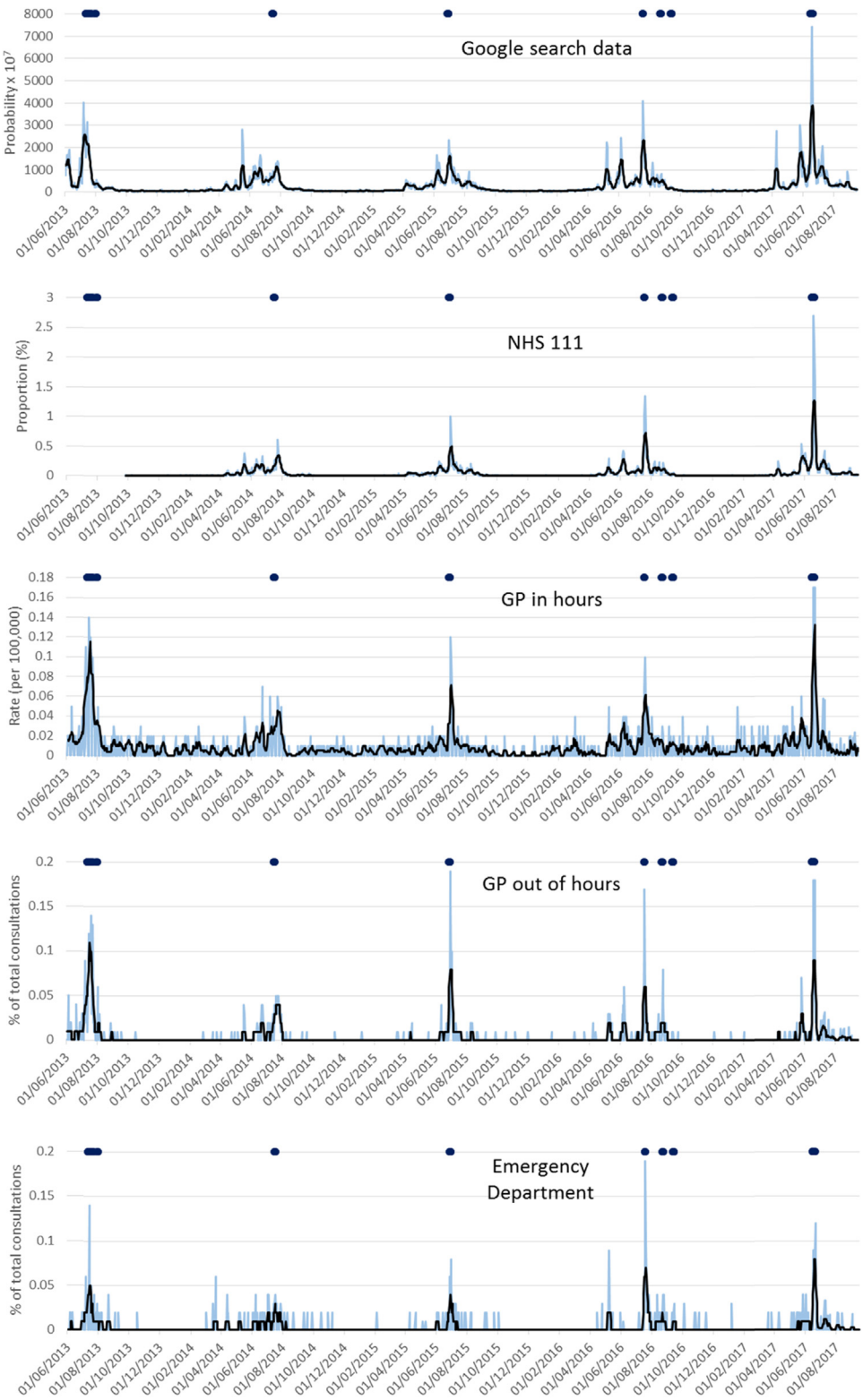

Fig. 1. Daily (pale blue) and seven day moving average (black) probabilities[1] searching for a heat-related term[2] in Google, and syndromic surveillance heat-related health indicators, England, 2013–2017[3]. NHS 111 = a national telehealth service, GP in hours = an in-hours general practitioner system, GP out of hours = an out of hours general practitioner system, Emergency Department = an emergency department surveillance system. [1]Daily probability is defined as searching for a specified term on a given day in England, multiplied by ten million. [2]Heat exhaustion or heat rash or heat stroke or heatwave or prickly heat or sunburn or sunstroke. [3]NHS 111 data was only available from September 2013 onwards. Periods with heatwave alerts greater than level 1 are indicated with blue dots (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

lagging or leading syndromic data respectively. Analysis was restricted to national level according to Google data availability.

PHE has approval under Section 251 of the NHS Act 2006 to process confidential patient information for public health purposes; separate ethical approval was not required for this study.

3. Results

The key Google search terms identified were *heat exhaustion*, *prickly heat*, *heat rash*, *heat stroke*, *heatwave*, *sunstroke* and *sunburn*. Each of the search terms showed an increase in frequency during periods of hot weather in each summer, with the largest increase seen in 2013 and 2017, corresponding with longer heatwave periods (Fig. 1 and

Table 2
Correlation coefficients and lag analysis between Google search term probability[4] and syndromic surveillance system heat-related indicator, England, 2013–2017[5].

Syndromic Surveillance system	Google Search Term	Spearman correlation coefficient			Lag analysis during 2017 heatwave										
		2013 to 2017	June 2017 Heatwave		Google dataset lags (days behind)[6]						Google dataset leads (days ahead)[7]				
			Daily	7dma[8]	-5	-4	-3	-2	-1	0	1	2	3	4	5
111	Heat exhaustion	**0.754**	**0.721**	**0.918**	0.097	0.331	**0.571**	**0.680**	**0.736**	**0.721**	**0.725**	**0.567**	0.365	0.128	-0.162
GPIH		**0.526**	0.280	**0.488**	-0.078	0.060	0.372	**0.392**	0.286	0.280	0.322	0.285	0.271	0.274	0.151
GPOOH		**0.628**	**0.673**	**0.855**	0.126	0.341	0.326	**0.560**	**0.599**	**0.673**	**0.704**	**0.538**	**0.432**	0.215	0.021
ED		**0.548**	**0.531**	**0.863**	-0.002	0.187	0.247	**0.503**	**0.593**	**0.531**	**0.514**	**0.566**	**0.484**	0.263	-0.056
111	Heat rash	**0.856**	**0.828**	**0.928**	0.132	0.358	**0.570**	**0.708**	**0.800**	**0.828**	**0.770**	**0.573**	0.293	-0.010	-0.243
GPIH		**0.480**	0.218	**0.480**	0.158	0.232	0.230	0.336	0.317	0.218	0.345	**0.393**	0.214	0.110	0.106
GPOOH		**0.643**	**0.758**	**0.874**	0.130	0.280	**0.472**	**0.656**	**0.718**	**0.758**	**0.617**	**0.467**	0.348	0.105	-0.145
ED		**0.533**	**0.594**	**0.833**	0.149	0.293	0.330	**0.464**	**0.572**	**0.594**	**0.523**	**0.412**	0.327	0.144	-0.061
111	Heat stroke	**0.684**	**0.916**	**0.958**	-0.082	0.122	**0.413**	**0.616**	**0.779**	**0.916**	**0.873**	**0.757**	**0.535**	0.285	0.008
GPIH		**0.410**	0.336	**0.482**	0.001	0.021	0.207	0.283	0.240	0.336	**0.405**	0.328	0.228	0.279	0.156
GPOOH		**0.574**	**0.680**	**0.923**	0.037	0.205	**0.394**	**0.530**	**0.709**	**0.680**	**0.684**	**0.607**	**0.552**	0.266	0.113
ED		**0.476**	**0.537**	**0.819**	-0.030	0.184	0.221	0.335	**0.515**	**0.537**	**0.527**	**0.498**	**0.497**	0.335	0.214
111	Heatwave	**0.774**	**0.450**	**0.667**	0.043	0.266	0.387	**0.461**	**0.515**	**0.450**	**0.480**	**0.394**	0.289	0.218	0.169
GPIH		**0.458**	**0.396**	**0.427**	-0.100	0.033	0.112	0.161	0.311	**0.396**	**0.409**	0.351	0.310	0.172	0.075
GPOOH		**0.514**	**0.522**	**0.704**	0.133	0.324	0.293	**0.454**	**0.514**	**0.522**	**0.534**	**0.609**	**0.423**	**0.450**	0.228
ED		**0.497**	**0.575**	**0.625**	-0.101	0.062	0.215	0.267	**0.452**	**0.575**	**0.518**	**0.527**	0.339	0.286	0.051
111	Prickly heat	**0.900**	**0.812**	**0.929**	0.168	0.358	**0.643**	**0.749**	**0.778**	**0.812**	**0.729**	**0.562**	0.276	0.048	-0.182
GPIH		**0.504**	0.344	**0.482**	0.167	0.014	0.294	**0.404**	0.244	0.344	**0.414**	0.301	0.137	0.192	0.118
GPOOH		**0.638**	**0.642**	**0.825**	0.200	0.304	**0.509**	**0.534**	**0.730**	**0.642**	**0.595**	**0.534**	0.355	0.179	-0.146
ED		**0.531**	**0.652**	**0.869**	0.092	0.219	0.359	0.347	**0.587**	**0.652**	**0.565**	**0.450**	0.301	0.157	-0.008
111	Sunburn	**0.898**	**0.844**	**0.956**	0.095	0.201	**0.421**	**0.560**	**0.722**	**0.844**	**0.844**	**0.696**	**0.402**	0.108	-0.162
GPIH		**0.506**	0.137	**0.493**	0.232	0.126	0.253	0.313	0.182	0.137	**0.374**	**0.444**	0.215	0.248	0.155
GPOOH		**0.629**	**0.693**	**0.908**	0.097	0.276	0.421	**0.492**	**0.691**	**0.693**	**0.618**	**0.489**	**0.402**	0.071	-0.129
ED		**0.528**	**0.563**	**0.835**	0.135	0.283	0.297	0.321	**0.501**	**0.563**	**0.493**	**0.375**	0.344	0.189	0.015
111	Sunstroke	**0.884**	**0.811**	**0.943**	-0.143	-0.012	0.264	**0.452**	**0.653**	**0.811**	**0.869**	**0.776**	**0.556**	0.325	0.069
GPIH		**0.476**	0.109	**0.464**	0.156	-0.063	0.079	0.219	0.118	0.109	**0.448**	**0.480**	0.161	0.200	0.236
GPOOH		**0.634**	**0.655**	**0.902**	-0.105	0.073	0.243	**0.394**	**0.645**	**0.655**	**0.555**	**0.524**	**0.575**	0.322	0.049
ED		**0.547**	**0.555**	**0.825**	-0.084	0.123	0.170	0.212	**0.445**	**0.555**	**0.458**	**0.395**	**0.467**	0.345	0.161

111 = a national telehealth service, GPIH = an in-hours general practitioner system, GPOOH = an out of hours general practitioner system, ED = an emergency department surveillance system.
[4]Daily probability is defined as searching for a specified term on a given day in England, multiplied by ten million. [5]Correlation coefficients are emboldened and shaded pale yellow where significant ($p < 0.05$). Cells shaded yellow contain the highest correlation coefficient for each row in the lag analysis. [6]The Google dataset was lagged backwards in time. E.g. The highest correlation coefficient with a lag of − 1 implies Google data needed to be shifted back a day to best correlate with syndromic data and it had a delay of one day relative to syndromic data. [7]The Google dataset was lagged forwards in time. E.g. The highest correlation coefficient with a lag of + 1 implies Google data needed to be shifted forwards a day to best correlate with syndromic data and it was one day ahead of syndromic data. [8]7dma = seven day moving average.

Supplementary Figure 1). In 2017, the most frequently searched term out of this list was *sunburn* (peak probability on 18th June), followed by *heatwave* (18th June), *heat stroke* (19th June), *heat rash* (20th June), *sunstroke* (18th June), *prickly heat* (19th June) and *heat exhaustion* (19th June). These increases coincided with heatwave alerts (Supplementary Table 1).

There was a positive correlation observed between daily Google search term and syndromic data (Fig. 1). Correlation coefficients between search term frequency and syndromic indicators from 2013 to 2017 for the different systems demonstrated the highest correlation with NHS 111 (range of 0.684–0.900 by search term) and the lowest with GPIH (range of 0.410–0.526, Table 2). When focusing on the heatwave event in June 2017, correlation coefficients increased and all were statistically significant except for GPIH daily data which was only significantly correlated with *heatwave*.

Lag analysis demonstrated maximal correlation between the Google and syndromic data within two days of each other across search terms (Table 2). By surveillance system, for most search terms GPIH was preceded by Google data by a day or two. There was no consistent pattern by search term across surveillance systems.

4. Discussion

Having national real-time surveillance systems in place to rapidly detect on a daily basis the health impact of heatwave events affords a timelier response and the information can also be used for the planning of interventions to minimise future impact (PHE 2015). This area is of increasing importance with the implications of climate change and predicted increases in temperature likely leading to more frequent and severe heatwaves (Hajat et al., 2014).

To the best of our knowledge, this is the first study to look at the daily frequency of Google search terms indicative of heat-related morbidity and compare these to indicators from established syndromic surveillance systems. There were increases in frequency across search terms nationally in England during heatwave periods from 2013 to 2017.

These increases correlated well with validated syndromic indicators of heat-related morbidity from NHS 111, and correlations improved for other established syndromic surveillance systems when focusing on the heatwave event in 2017. The higher degree of correlation with the NHS 111 telehealth service could indicate people are searching for symptoms corresponding to the less severe conditions individuals would typically report to a telehealth service, e.g. sunburn, and the individuals who use NHS 111 are more likely to have searched on the internet as well. Conversely, the lower correlation with GP and ED consultation data could indicate people using Google a few days later to get more information about their diagnosis, supported by lag analysis observations.

Known limitations of Google search data include the representativeness of internet users. The most recent data suggest 89% of adults in the UK had recently used the internet and while use in the elderly is lower (40% in 75 + year olds), it is starting to increase (ONS, 2017). While digital inequalities in internet use have been seen (Reisdorf and Groselj, 2014), there is still a lot of uncertainty around internet use amongst vulnerable groups such as the homeless, and amongst culturally and linguistically diverse communities. However, despite these

limitations search engine market share data indicate the majority of all searches (85%) are carried out in Google and so this is an appropriate search engine to use in such studies (Statista, 2018).

Other limitations include the inability to look at subnational data to correspond to the regional heatwave alerts and uncertainty around who is searching for information and the intention behind the search, including the possibility of searches driven by the media. Therefore, it may not be as indicative of poor health as existing syndromic data sources, which are able to provide subnational data and information on the characteristics of the person such as age and sex. However, the correlation with the timing of syndromic indicators, the consistency of the pattern across a range of search terms and the selection of symptom-based terms suggest people were searching for information because of a health complaint (either for themselves or someone else). Furthermore, associated search terms identified in Google Trends for all queries except for *heatwave* were *signs*, *symptoms* and *treatment*. For *heatwave*, associated search terms focused on the specific event such as *UK*, *weather* and *2016*, suggesting the term was used to search for information about the event rather than specifically for health effects during the event. Exploratory qualitative research with a representative sample of Google users will help to determine the reasoning behind searching Google for such terms and validate if searched for because of a health complaint.

Lag analysis revealed a similar timeliness between daily Google and syndromic datasets during the heatwave in June 2017, suggesting Google data provided no time advantage over syndromic data (i.e. did not provide a delayed or earlier signal). The exception to this was GPIH where Google data provided a slightly earlier signal of one or two days. This may be explained by the unavailability of routine GP services during weekends leading to a delay in patients consulting. In the case of June 2017, the peak of the heatwave occurred during a weekend.

These findings have highlighted two key potential benefits. Firstly, Google-based syndromic surveillance will be of interest to those countries lacking established public health surveillance systems able to monitor the health impact of heat in near real-time but could access freely available Google data. Secondly, this could contribute to assessing the wider population health impact of exposure to heat. The indicators monitored through syndromic surveillance systems are focussed on health outcomes severe enough to require healthcare to provide near real-time information on whether the severity of the impact of a heatwave is higher than expected, typically in conjunction with excess mortality surveillance (Green et al., 2016). Monitoring the frequency of internet search queries could complement existing surveillance of heat-related morbidity by supporting the estimation of the proportion of the population whose health may be affected by heat who may not present to healthcare services.

While expanding on previous work that looked at single search terms, a limited number of search terms in English were individually assessed in this study, with a focus on heat outcomes. Further consideration should be given to looking at trends in Google search term frequency for specific health impacts of heatwaves such as heart disease, foodborne disease and work-related injuries. Further modelling work will be required to both fully develop a more specific and validated search term list for Google using natural language processing methods (Lampos et al., 2017) and to use these outputs in combination to estimate the wider population health impact of heatwaves.

This work forms part of a wider programme of work being carried out by PHE to assess the utility of Google search data for the surveillance of a range of health outcomes in the context of existing public health syndromic surveillance. The work presented here demonstrates the potential utility of internet search queries, particularly in countries where no such daily healthcare-based syndromic surveillance systems are in place. Search terms may complement existing heat health surveillance systems in England, with multiple rapid surveillance systems adding a level of confidence to monitoring. Further work is required to validate and interpret the findings prospectively.

Acknowledgements

We acknowledge support from: NHS 111 and NHS Digital for their assistance and support with the NHS 111 system; OOH providers submitting data to the GPOOHSS and Advanced Heath & Care; ED clinicians and NHS Trust staff supporting EDSSS, the Royal College of Emergency Medicine and technical support provided by EMIS Health and L2S2 Ltd; TPP and participating SystmOne practices and University of Nottingham, ClinRisk, EMIS and EMIS practices submitting data to the QSurveillance database and Google for access to their Google Health Trends API. The authors would like to thank Mr Paul Loveridge of the PHE Real-time Syndromic Surveillance Team (ReSST) and Dr Bill Lampos of the Department of Computer Science, UCL for their technical expertise. IJC and RAM are supported by the i-sense EPSRC IRC in Early Warning Sensing Systems for Infectious Diseases (EP/K031953/1) and RAM is also supported by the Royal Society Wolfson Merit Award. AJE and GES are supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emergency Preparedness and Response at King's College London in partnership with PHE. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

Funding

This work was undertaken as part of the national surveillance function of Public Health England and received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

None declared.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2018.04.002.

References

Benmarhnia, T., Deguen, S., Kaufman, J.S., Smargiassi, A., 2015. Review article: Vulnerability to heat-related mortality: a systematic review, meta-analysis, and meta-regression analysis. Epidemiology 26 (6), 781–793.

Bouchama, A., Dehbi, M., Mohamed, G., Matthies, F., Shoukri, M., Menne, B., 2007. Prognostic factors in heat wave related deaths: a meta-analysis. Arch. Intern. Med. 167 (20), 2170–2176.

Buckingham-Jeffery, E., Morbey, R., House, T., Elliot, A.J., Harcourt, S., Smith, G.E., 2017. Correcting for day of the week and public holiday effects: improving a national daily syndromic surveillance service for detecting public health threats. BMC Public Health 17, 477.

Elliot, A.J., Bone, A., Morbey, R., Hughes, H.E., Harcourt, S., Smith, S., Loveridge, P., Green, H.K., Pebody, R., Andrews, N., Murray, V., Catchpole, M., Bickler, G., McCloskey, B., Smith, G., 2014. Using real-time syndromic surveillance to assess the health impact of the 2013 heatwave in England. Environ. Res. 135, 31–36.

Green, H.K., Andrews, N., Armstrong, B., Bickler, G., Pebody, R., 2016. Mortality during the 2013 heatwave in England–How did it compare to previous heatwaves? A retrospective observational study. Environ. Res. 147, 343–349.

Google, 2017. Google Trends. Available online: ⟨https://trends.google.com/trends/⟩.

Hajat, S., Vardoulakis, S., Heaviside, C., Eggen, B., 2014. Climate change effects on human health: projections of temperature-related mortality for the UK during the 2020s, 2050s and 2080s. J. Epidemiol. Community Health 68 (7), 641–648.

Harcourt, S.E., Morbey, R.A., Loveridge, P., Carrilho, L., Baynham, D., Povey, E., Fox, P., Rutter, J., Moores, P., Tiffen, J., Bellerby, S., McIntosh, P., Large, S., McMenamin, J., Reynolds, A., Ibbotson, S., Smith, G.E., Elliot, A.J., 2017. Developing and validating a new national remote health advice syndromic surveillance system in England. J. Public Health 39 (1), 184–192.

Josseran, L., Fouillet, A., Caillère, N., Brun-Ney, D., Ilef, D., Brucker, G., Medeiros, H., Astagneau, P., 2010. Assessment of a syndromic surveillance system based on morbidity data: results from the Oscour network during a heat wave. PLoS One 5 (8), e11984.

Lall, R., Abdelnabi, J., Ngai, S., Parton, H.B., Saunders, K., Sell, J., Wahnich, A., Weiss, D., Mathes, R.W., 2017. Advancing the use of emergency department syndromic surveillance data, New York City, 2012–2016. Public Health Rep. 132 (1_suppl),

23S–30S.
Lampos V., Zou B., Cox I.J., 2017. Enhancing Feature Selection Using Word Embeddings: The Case of Flu Surveillance. Proceedings of the 26th International Conference on World Wide Web. In: International World Wide Web Conferences Steering Committee. p. 695–704.
Li, M., Gu, S., Bi, P., Yang, J., Liu, Q., 2015. Heat waves and morbidity: current knowledge and further direction-a comprehensive literature review. Int. J. Environ. Res. Public Health 12 (5), 5256–5283.
Li, T., Ding, F., Sun, Q., Zhang, Y., Kinney, P.L., 2016. Heat stroke internet searches can be a new heatwave health warning surveillance indicator. Sci. Rep. 6, 37294.
Met Office, 2017. Heat-health watch. Available online: ⟨http://www.metoffice.gov.uk/public/weather/heat-health/#?Tab=heatHealth⟩.
Nelson, N.G., Collins, C.L., Comstock, R.D., McKenzie, L.B., 2011. Exertional heat-related injuries treated in emergency departments in the U.S. 1997–2006. Am. J. Prev. Med. 40 (1), 54–60.
Nuti, S.V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R.P., Chen, S.I., Murugiah, K., 2014. The use of google trends in health care research: a systematic review. PLoS One 9 (10), e109583.
Office for National Statistics, 2017. Internet users in the UK: 2017. Available online: ⟨https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2017⟩.
Pascal, M., Laaidi, K., Wagner, V., Ung, A.B., Smaili, S., Fouillet, A., Caserio-Schönemann, C., Beaudeau, P., 2012. How to use near real-time health indicators to support decision-making during a heat wave: the example of the French heat wave warning system. PLoS Curr. 4 (e4f83ebf72317d).
Public Health England, 2017. Syndromic surveillance: systems and analyses. Available online: ⟨https://www.gov.uk/government/collections/syndromic-surveillance-systems-and-analyses⟩.
Public Health England, 2015. Heatwave plan for England. Protecting health and reducing harm from severe heat and heatwaves. Available online: ⟨https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/429384/Heatwave_Main_Plan_2015.pdf⟩.
Ramin, B., Tomislav, S., 2009. Health of the homeless and climate change. J. Urban Health 86 (4), 654–664.
Reisdorf B., Groselj D., 2014. For what it's worth: digital inequalities, attitudes and a typology of internet (non-) users. In Proceedings of the 2014 ACM conference on Web science pp. 263–264.
Robine, J.M., Cheung, S.L., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.P., Herrmann, F.R., 2008. Death toll exceeded 70,000 in Europe during the summer of 2003. C. R. Biol. 331 (2), 171–178.
Smith, S., Elliot, A.J., Hajat, S., Bone, A., Smith, G.E., Kovats, S., 2016a. Estimating the burden of heat illness in England during the 2013 summer heatwave using syndromic surveillance. J. Epidemiol. Community Health 70 (5), 459–465.
Smith, S., Elliot, A.J., Hajat, S., Bone, A., Bates, C., Smith, G.E., Kovats, S., 2016b. The impact of heatwaves on community morbidity and healthcare usage: a retrospective observational study using real-time syndromic surveillance. Int. J. Environ. Res Public Health 13 (1).
Statista, 2018. Market share held by the leading search engines in the United Kingdom (UK) as of January 2018. Available online: ⟨https://www.statista.com/statistics/280269/market-share-held-by-search-engines-in-the-united-kingdom/⟩.
Triple S Project, 2011. Assessment of syndromic surveillance in Europe. Lancet 378 (9806), 1833–1834.